Energy Price Forecasting with Uncertainty Estimation

Master Thesis Presentation by

Sneha Senthil

Chair of Algorithms and Data Structures
University of Freiburg

CONTENTS

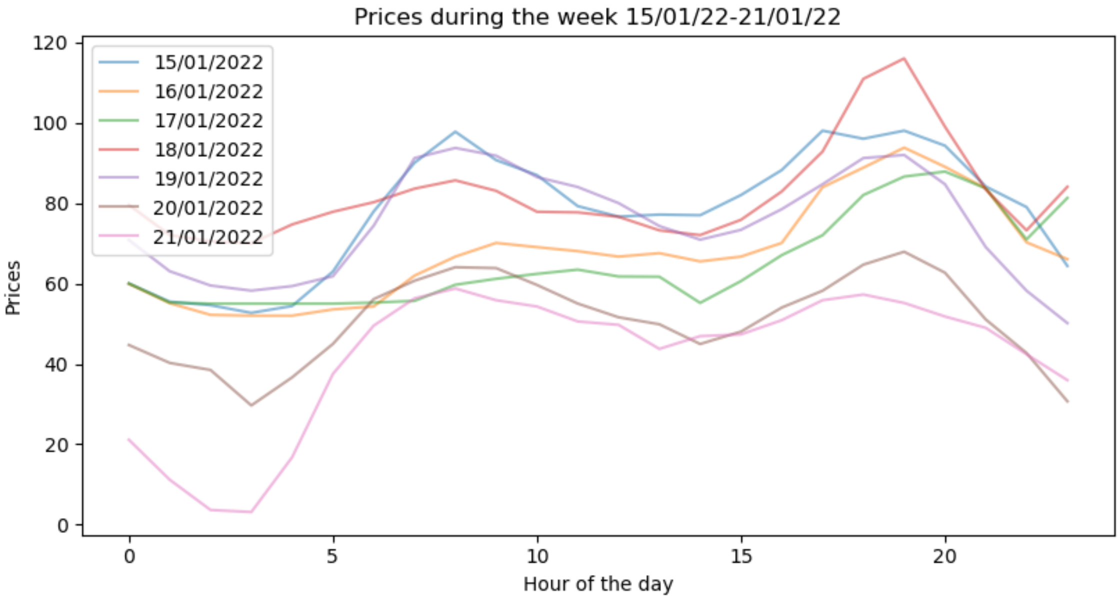
UNI FREIBURG

INTRODUCTION

- Electricity is bought in the day-ahead market.
- Balancing supply versus demand leads to highly volatile market prices.
- Predicting the day-ahead price would help maximize profit.

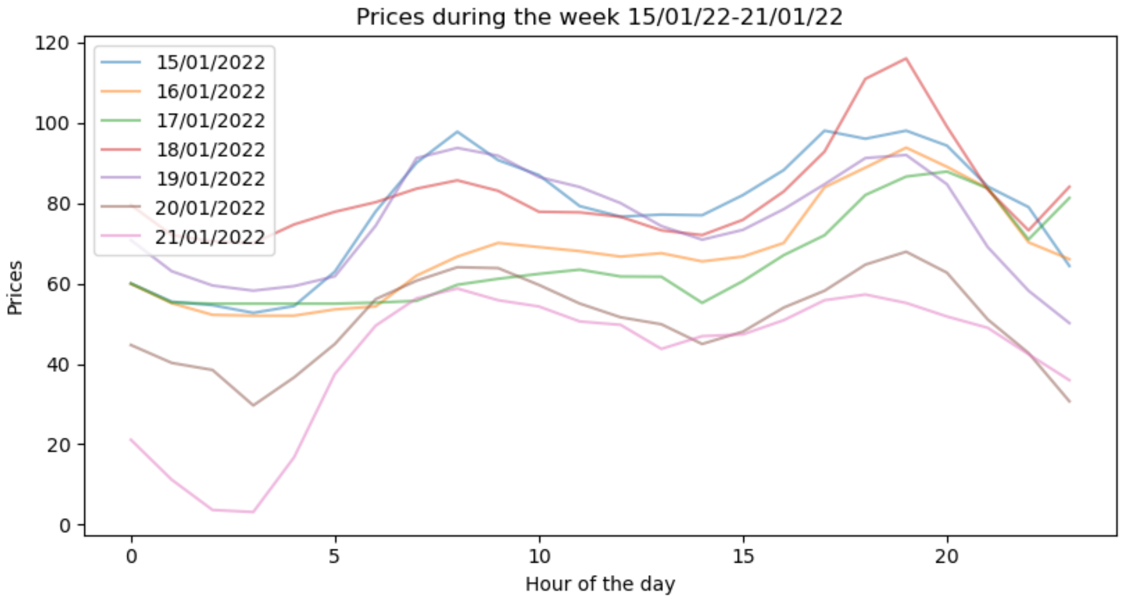
UNI FREIBURG

MOTIVATION



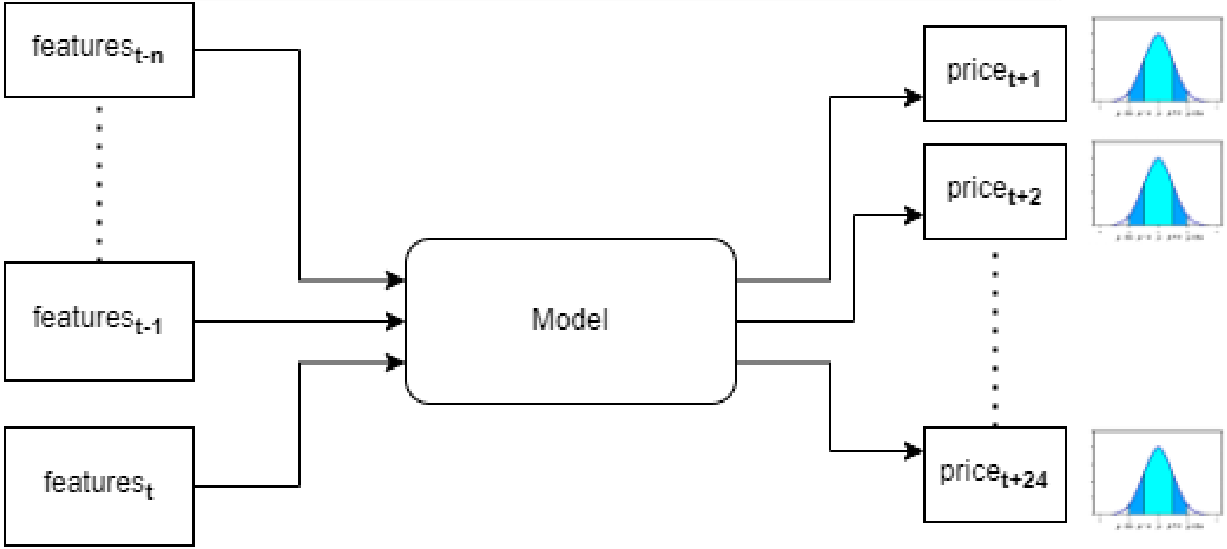
UNI FREIBURG

MOTIVATION



Prices during the week 15/01/22-21/01/22
15/01/2022
16/01/2022
17/01/2022
18/01/2022
19/01/2022
20/01/2022
21/01/2022
Prices
Hour of the day
0
20
40
60
80
100
120
0
5
10
15
20

PROBLEM



features are load, prices, wind energy, solar energy and weather features

Image retrieved from http://www.stat.yale.edu/Courses/1997-98/101/normal.htm

INTRODUCTION

Questions?

CONTENTS

1. Introduction
2. **Methods**
3. Evaluation

Approach

Download the data
Combine datasets
Data pre-processing
Create training, validation and test sets
Split data into corresponding features and labels
Train models
Evaluate and compare

METHODS

- Data
- Models

DATA

Data for 2 countries is considered: **Spain** and **Switzerland**. Models are trained separately on each dataset.

Features from Entsoe:

- Load
- Generation (Solar and Wind)
- Prices

Features from Copernicus:

- Shortwave Radiation
- Wind Speed
- Air Temperature
- Total Precipitation

DATA- FEATURES AND LABELS

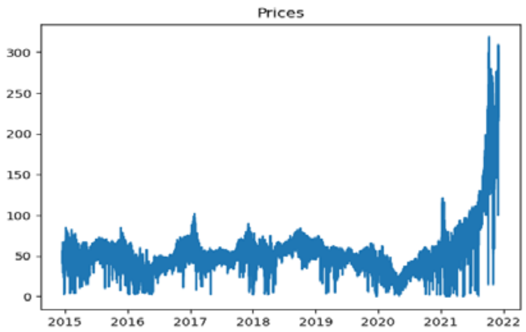
<u>Features</u>: Hourly historical features, going back n-hours

	f1	f2	f3	f4	f5	f6	f7	f8
t-n								
....								
t-2								
t-1								
t								

t+1
t+2
t+3
……
t+24

<u>Labels</u>: Next-day prices (24 hours)

DATA AUGMENTATION



Electricity Prices in Spain

Electricity Prices in Switzerland

Data augmentation is done by multiplying the training data with 3.5. This is used as additional training data.

UNI FREIBURG

MODELS

- Deterministic Models
- Probabilistic Models

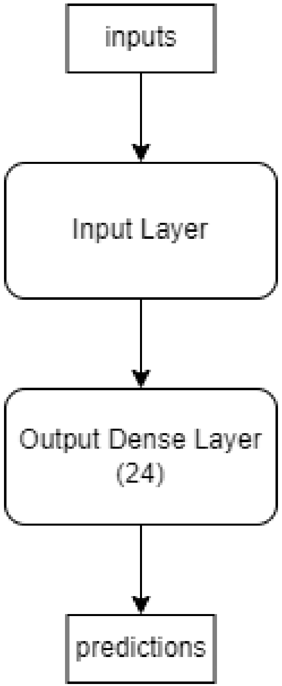
DETERMINISTIC MODELS

- Linear model (Baseline)
- Residual MLP
- LSTM
- Transformer

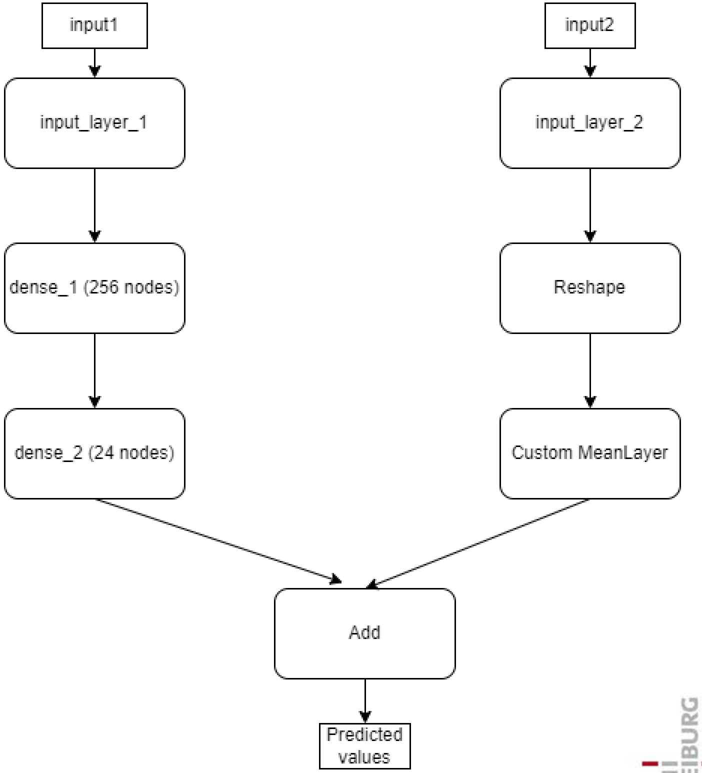
  -The first 3 models were used during the master project. They are now being used to compare against the Transformer model.

DETERMINISTIC MODELS- LINEAR

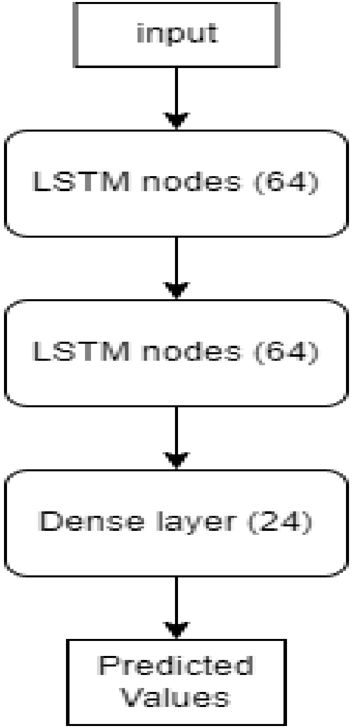
- Used as a baseline



DETERMINISTIC MODELS- RESIDUAL MLP



DETERMINISTIC MODELS- LSTM



DETERMINISTIC MODELS- Transformer

- State of the art model for many NLP tasks

- We use a transformer encoder-only model for forecasting energy prices

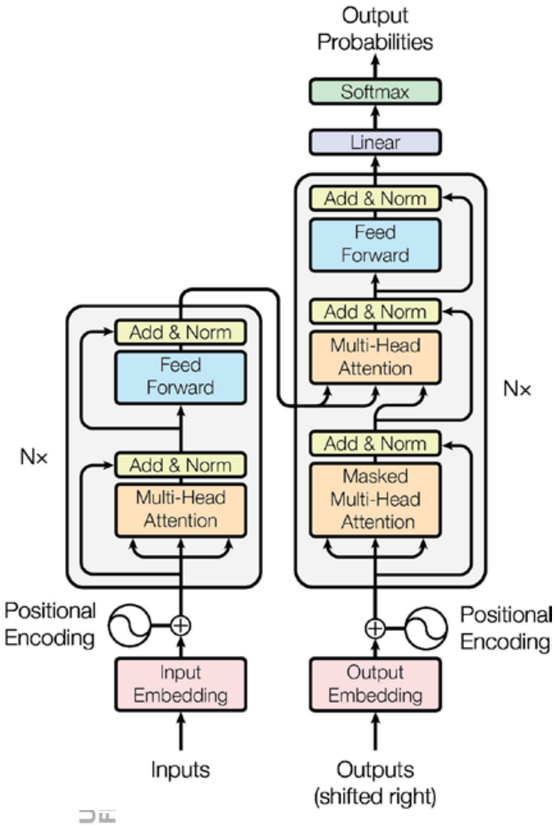
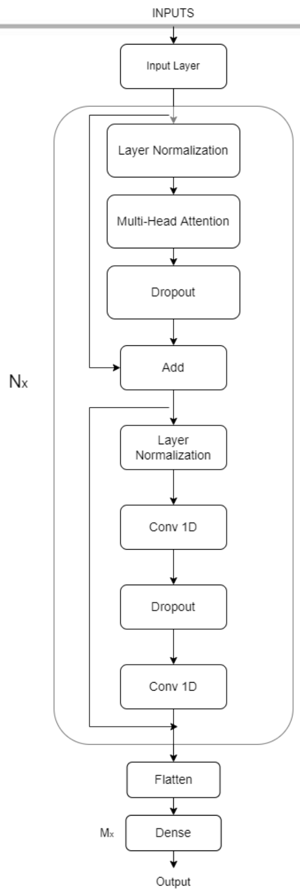


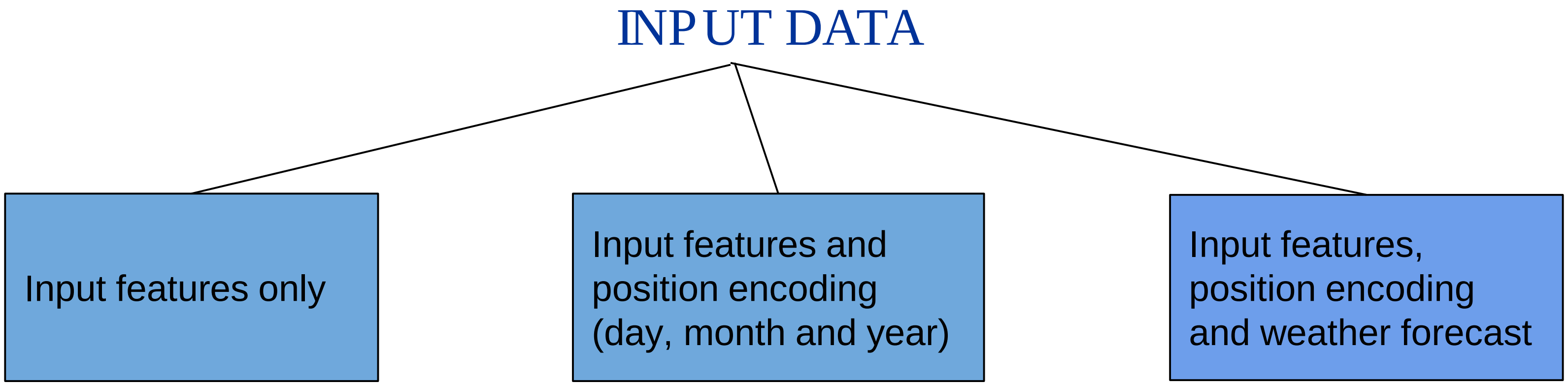
Image retrieved from Vaswani et al. (2017)

DETERMINISTIC MODELS- Transformer



DETERMINISTIC MODELS- Transformer

Based on the input features to the model, there are 3 types of Transformer models that were trained.



Hyperparameters

HYPERPARAMETER	VALUE
number of hours back	72
number of hours back (Transformer)	168
Initial learning rate	10^{-4}
Optimizer	Adam

Regularisation: Early Stopping, Reduce Learning Rate on Plateau, Learning Rate Scheduler

UNI FREIBURG

DETERMINISTIC MODELS- RECAP

There are 4 deterministic models used for energy price prediction:

1. Linear Model
2. Residual MLP
3. LSTM
4. Transformer (3 different models based on the input data)

UNI FREIBURG

MODELS

- Deterministic Models
- **Probabilistic Models**

PROBABILISTIC MODEL

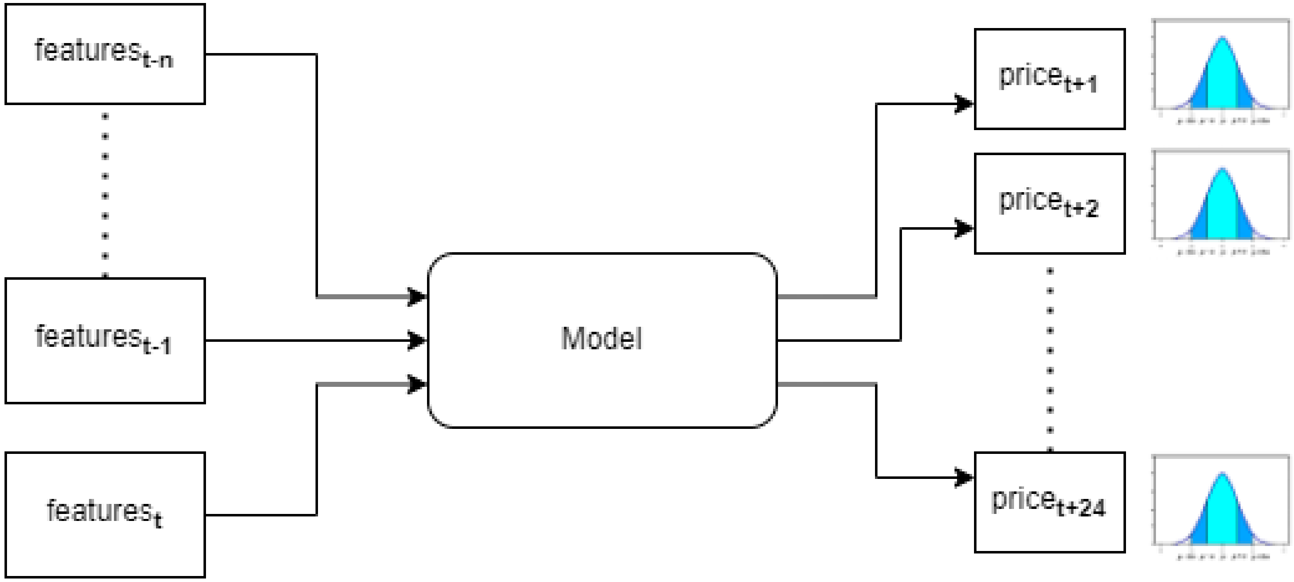
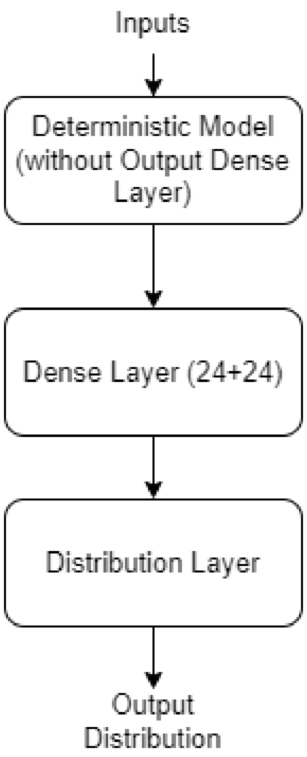


Image retrieved from http://www.stat.yale.edu/Courses/1997-98/101/normal.htm

PROBABILISTIC MODEL



PROBABILISTIC MODEL

3 distributions were considered while training the probabilistic models:

Normal Distribution

Learnable parameters:
- Mean
- Standard Deviation

Lognormal Distribution

Learnable parameters:
- Mean
- Standard Deviation

Gamma Distribution

Learnable parameters:
- Concentration
- Rate

METHODS

Questions?

CONTENTS

UNI FREIBURG

EVALUATION

- Deterministic Models
- Probabilistic Models

EVALUATION- DETERMINISTIC MODELS

Evaluation Metrics:

- Mean Absolute Error

$$MAE = \frac{1}{n}\Sigma_{j=1}^{n}|y_j - \hat{y}_j|$$

- Root Mean Squared Error

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

UNI FREIBURG

EVALUATION- DETERMINISTIC MODELS

RESULTS - Spain Dataset

MODEL	MAE (€/MWh)	RMSE (€/MWh)
Linear (Baseline)	47.98	59.63
Residual MLP	13.42	19.17
LSTM	13.18	19.22
Transformer	9.75	15.09

Mean price in test data: 135.3 €/MWh

RESULTS - Switzerland Dataset

MODEL	MAE (€/MWh)	RMSE (€/MWh)
Linear (Baseline)	85.9	106.92
Residual MLP	22.64	32.18
LSTM	21.71	31.4
Transformer	17.29	26.05

Mean price in test data: 164.5 €/MWh

UNI FREIBURG

EVALUATION- TRANSFORMER

RESULTS - Input Features Dependency (Spain Dataset)

INPUT DATA	MAE (€/MWh)	RMSE (€/MWh)
Only Features	9.75	15.09
Features+Position Encoding	11.6	16.75
Features+Position Encoding+Weather Forecast	11.64	17.02

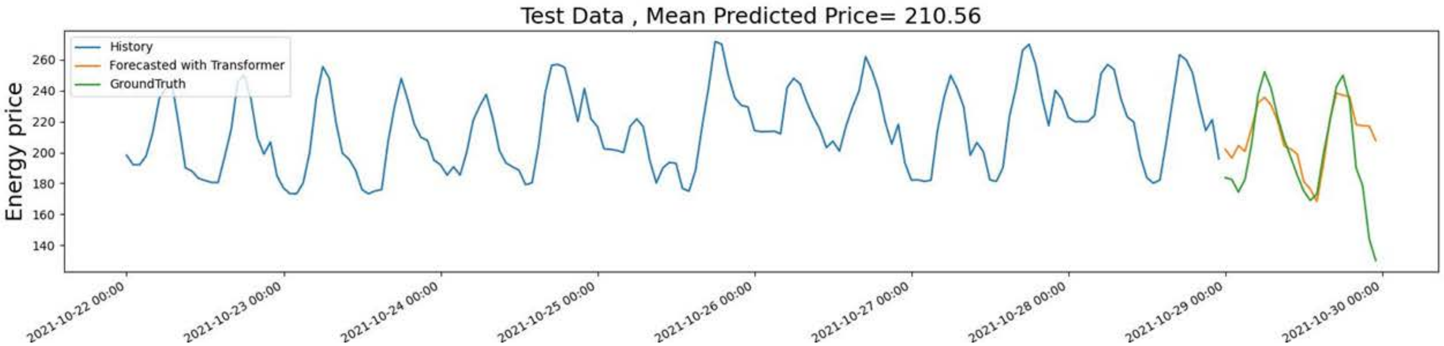
RESULTS - Input Features Dependency (Switzerland Dataset)

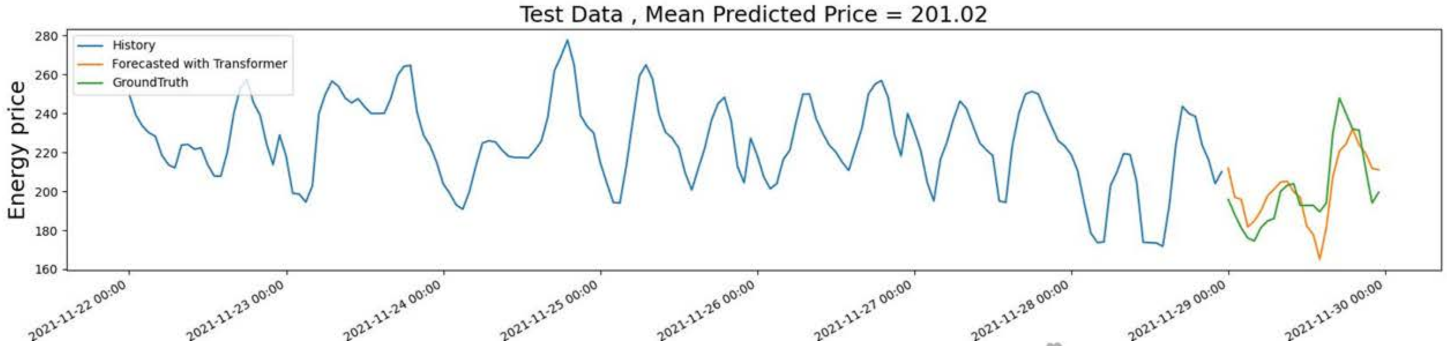
INPUT DATA	MAE (€/MWh)	RMSE (€/MWh)
Only Features	17.29	26.05
Features+Position Encoding	17.91	27.07
Features+Position Encoding+Weather Forecast	19.1	28.45

EVALUATION- TRAINING TIMES

MODEL	TRAINING TIME
Linear	4 minutes and 20 seconds
Residual MLP	27 minutes and 5 seconds
LSTM	150 minutes and 25 seconds
Transformer	12 minutes and 30 seconds

TRANSFORMER- PREDICTION GRAPHS





UNI FREIBURG

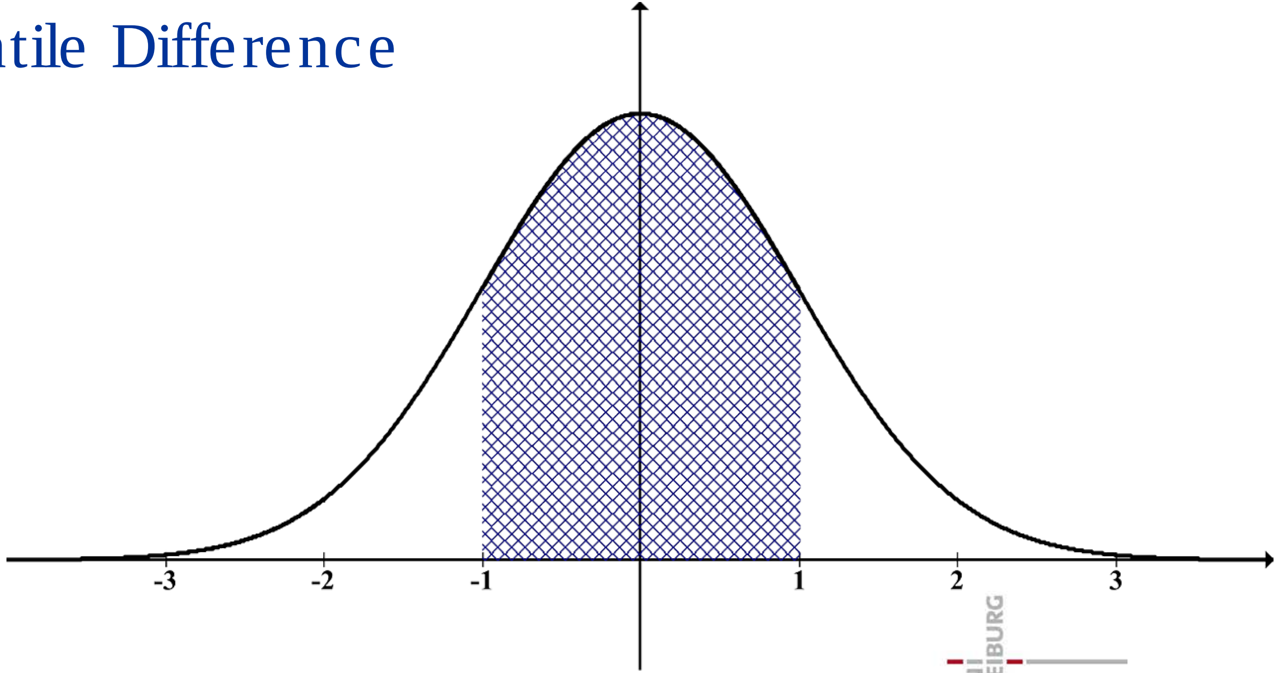
EVALUATION

- Deterministic Models
- **Probabilistic Models**

EVALUATION- PROBABILISTIC MODELS

Evaluation Metrics:

1) Quantile Difference



EVALUATION- PROBABILISTIC MODELS

Evaluation Metrics:

2) CRPS Score

- Continuous Ranked Probability Score
- It measures the squared distance between the predicted distribution and the target.

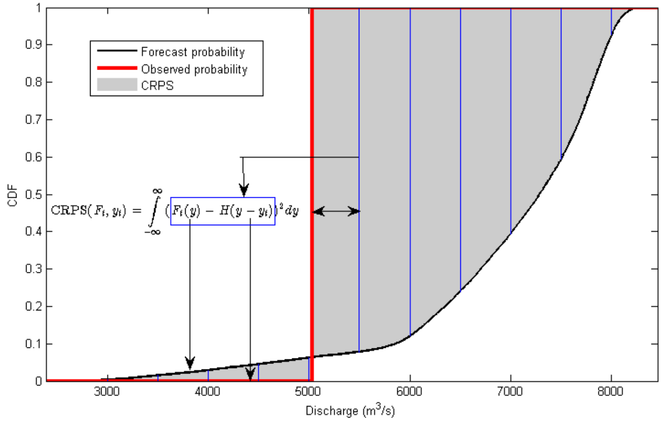


Image retrieved from https://www.mathworks.com/matlabcentral/fileexchange/47807-continuous-rank-probability-score

EVALUATION- PROBABILISTIC MODELS

Evaluation Metrics:

3) Log Likelihood

- Logarithm of the probability density function of the observed data
- The higher log likelihood value, the better the model is at fitting the data

EVALUATION- PROBABILISTIC MODELS

Results- Residual MLP (Spain)

DISTRIBUTION	**Quantile Difference (80%)**	**CRPS**	**Log Likelihood**	**MAE**
Normal	50.49	8.44	0.67	8.76
Log-normal	48.81	10.92	-0.67	10.17
Gamma	63.96	12.46	0.68	9.55

Deterministic Residual MLP: MAE = 13.42

UNI FREIBURG

EVALUATION- PROBABILISTIC MODELS

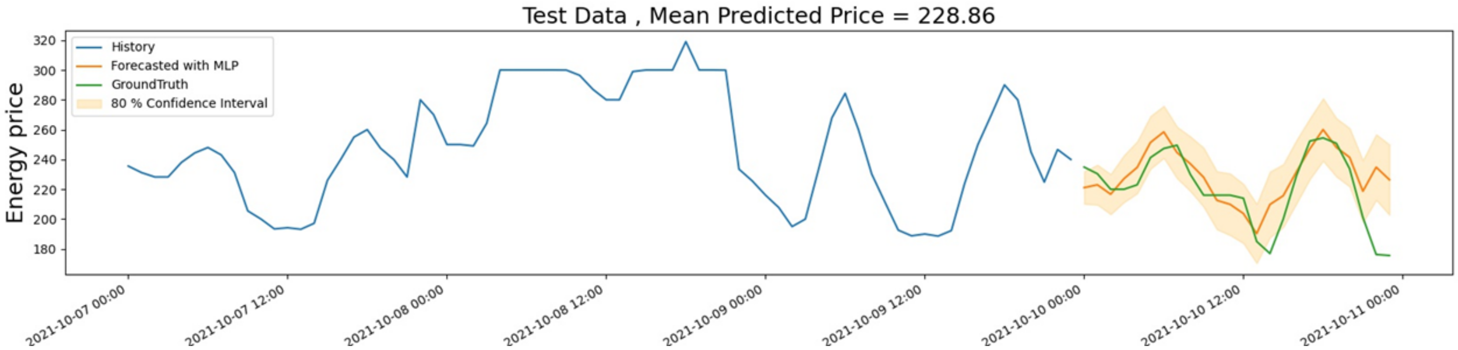
Results- Residual MLP (Spain)

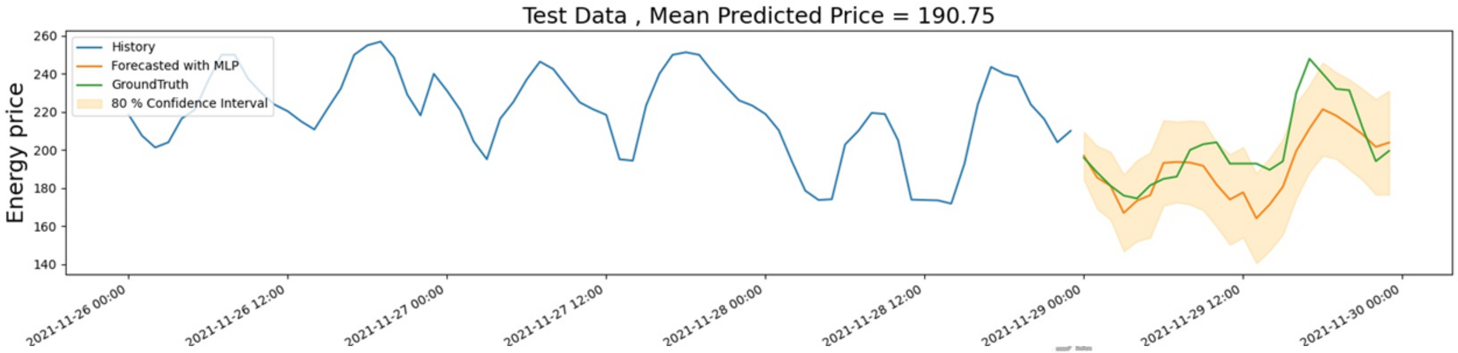
DISTRIBUTION	Quantile Difference (80%)	CRPS	Log Likelihood	MAE
Normal	50.49	8.44	0.67	10.85
Log-normal	48.81	10.92	-0.67	12.44
Gamma	63.96	12.46	0.68	12.47
Normal (Transformer)	48.81	12.39	0.27	15.96

Deterministic Transformer: MAE = 9.75

UNI FREIBURG

RESIDUAL MLP (NORMAL DISTRIBUTION)- PREDICTION GRAPHS





CONCLUSION

- Transformer is successful in time series forecasting.
- Converting a Residual MLP to a probabilistic model helped improve prediction accuracy, but this was not the case for the Transformer.
- Probabilistic predictions help in understanding uncertainty of the model.

UNI FREIBURG

Thank You!